Validating Files Transferred to Mass Storage by Size and Number

Category: Archiving Data

It is a good practice to check if files are copied correctly to Mass Storage. The simplest and most light weight method to check files sent to Lou1 or Lou2 is to compare the number of files and the disk space between the original and the copy.

Example:

```
pfe20% du -sk mydatadir
353760  mydatadir
pfe20% find mydatadir | wc -l
51
pfe20% scp -rp mydatadir lou2:

lou2% du -sk --apparent-size mydatadir
353684  mydatadir
lou2% find mydatadir | wc -l
51
```

Here the sizes are close and the number of files matches exactly.

Note that in most cases the directory size will not match exactly. The **--apparent-size** option is necessary on the Lou systems because the data may reside on tape, not disk.

Please note that the easiest way to verify the integrity of actual file contents during transfers is to use the NAS-developed Shift tool for the transfer with the **--verify** option enabled. Shift will automatically checksum the data at the source and destination to detect corruption as part of the transfer. If corruption is detected, partial file transfers/checksums will be performed until the corruption is rectified. For example:

```
pfe20% shiftc --verify $HOME/foo lou:
lou% shiftc --verify /nobackup2a/username/foo $HOME/foo
your_localhost% sup shiftc --verify foo lou:
```